



A Union of Professionals

American Federation of Teachers, AFL-CIO

WHAT'S PROFICIENT?

THE NO CHILD LEFT BEHIND ACT AND THE MANY MEANINGS OF PROFICIENCY

Bella Rosenberg
May 2004



WHAT'S PROFICIENT?

The No Child Left Behind Act and the Many Meanings of Proficiency

“Each State shall establish a timeline for adequate yearly progress. The timeline shall ensure that not later than 12 years after the end of the 2001-2002 school year, all students...will meet or exceed the State’s *proficient* level of academic achievements on the State assessments...” *No Child Left Behind Act of 2001* (emphasis added).

pro·fi·cient *adj.* Having or marked by an advanced degree of competence, as in an art, vocation, profession, or branch of learning. *n.* An expert; an adept. *The American Heritage® Dictionary of the English Language, Fourth Edition 2000.*

Table of Contents

Introduction.....	3
Different methods of setting performance standards (or achievement levels) yield different meanings of proficient.....	4
Different tests, even in the same subject, yield different percentages of proficient students.....	4
Proficient typically means different things in different grades, even on the same test and in the same subject.....	5
Different percentages of proficient students in different subjects do not necessarily mean that students or schools are stronger or weaker in one subject than the other.....	6
Proficient means different things in different states.....	7
NAEP proficiency standards are set very high.....	9
NCLB’s requirement of 100 percent proficiency is not the same thing as requiring that all students achieve at grade level.....	10

WHAT'S PROFICIENT?

The No Child Left Behind Act and the Many Meanings of Proficiency

The goal of the No Child Left Behind Act (NCLB) is to have 100 percent of America's public school students "proficient" by the year 2014, 12 years from the enactment of the law. Proficiency is to be measured through annual state-level tests in reading and math in each of grades 3-8 and at least once in high school.¹ In 2007-08, states must also test in science at least once in grades 3-5, 6-9, and 10-12.

The goal of 100 percent proficiency is certainly a giant leap forward from America's past expectations for "minimal competency" or universal attainment of "basic skills" or even being first in the world in international comparisons of student achievement. Indeed, NCLB's goal is globally unprecedented. Not even the highest-scoring nations that have participated in international tests of reading, math, and science are even close to attaining it.² Can 100 percent proficiency be accomplished?

Setting aside iron-clad facts about individual variability, the unprecedented, uniquely ambitious nature of NCLB's goal is not necessarily a reason to doubt that it can be realized by 2014. After all, Americans have set extremely challenging goals before – landing an astronaut on the moon comes to mind – and have attained them. And certainly 100 percent proficiency is at least as important, beneficial, and even thrilling a goal as landing an astronaut on the moon.

In the case of our moon goal, however, while most Americans didn't know what was involved in attaining it, we all knew what it meant and how we would know if it had been achieved: there would be an astronaut safely on the moon. But what does NCLB's goal of 100 percent proficiency mean? What's proficient?

Most parents and the public likely believe that 100 percent proficiency means that all students are on grade level. When U.S. Department of Education officials, as well as the media, say that the goal of NCLB and its "adequate yearly progress" (AYP) provision is having all students performing at grade level, they certainly encourage that belief.³ Are they right?

The chances are also good that if you asked the proverbial man on the street if a student who scored proficient on one fourth-grade math test would also score proficient on a different fourth-grade math test, he would answer yes. Moreover, if you asked him what his conclusion would be if School X had a higher percentage of students who were proficient in reading than in math, the likely answer would be that the school was doing a better job with reading. Similarly, if you told him that 50 percent of the fourth-graders in School Y were proficient in reading but only 37 percent of the fifth-graders were, he'd likely conclude that the fourth-grade teachers were clearly superior to the fifth-grade teachers at teaching the subject. True?

¹ States have until 2006 to test in these grades. Twenty states and the District of Columbia already do so. The rest may continue to test reading and math at least once between each of grades 4-6 and 6-8 and once in high school, as the previous version of the Elementary and Secondary Education Act required.

² Robert L. Linn, "Assessments and Accountability," *Educational Researcher*, 29, 2 (2000): pp. 4-16.

³ See, for example, the remarks of Eugene Hickok, acting deputy secretary of education, on the Lehrer NewsHour on March 15, 2004 (http://www.pbs.org/newshour/bb/education/jan-june04/schools_3-15.html); and Diana Jean Schemo, "14 States Ask U.S. To Revise Some Education Law Rules," *New York Times*, March 25, 2004, online edition.

And what about the inferences that are made about the quality of education when results from the National Assessment of Educational Progress (NAEP), the “nation’s report card,” come out?⁴ Who was not alarmed when, for example, former President Clinton reported that almost half of American fourth-graders were “below basic” in reading, or when presidential candidate George Bush ran ads in 2000 using similar NAEP results to say that almost half of our fourth-graders could barely handle a picture book? How many editorials decried our “failing” public schools when the most recent 12th-grade NAEP results showed “only” 36 percent of students at the proficient level in reading and an even paltrier 17 percent proficient in math?

But are all these perfectly understandable inferences that parents and the public typically make – and that politicians and the media so often promote – correct? Not necessarily. It all depends on how “proficient” is defined. Welcome, then, to the many meanings of “proficient” and a closer look at NCLB’s goal of 100 percent proficiency and making “adequate yearly progress” in getting there by 2014.

Different methods of setting performance standards (or achievement levels) yield different meanings of proficient.

“Proficient” is essentially a cut score (e.g., 90 out of 100 questions correct) on a test, though these days a “cut score” is often referred to as a performance standard or achievement level. There are a number of ways of deciding how to set that cut score, but under standards-based reform, it’s typically done by convening a committee, which usually includes many members from outside education, and putting them through one of many available “standard-setting” processes. These processes are more or less scientific, and each will yield a different result. Different committees using the same method may also come up with different results.

In the end, however, where cut scores are set under standards-based reform is a matter of *judgment* about how students *ought* to perform on a test. Some committees set the cut scores for different levels of achievement (e.g., basic, proficient, advanced) very high, others lower, and others in between. So, for example, even if two different states used the same test, a student in one state might have to score 90 to be considered proficient, while in the other, the student could be proficient with a 60. All things being equal, states that have higher cut scores for proficiency will have higher AYP failure rates than states with less challenging cut scores.

Different tests, even in the same subject, yield different percentages of proficient students.

Below is a table displaying the percentage of students labeled “proficient” or “advanced” on three different achievement test batteries from major test publishers and NAEP. Each of these national tests actually uses the label “proficient” and “advanced.” But note the widely varying results in the percentages of students who score proficient on them.

⁴ NAEP is administered nationally to a representative group of students in the fourth, eighth, and 12th grades and has been tracking progress in key academic subjects since 1968. In 1990, NAEP was also given the authority to report state-level results, but participation by the states was voluntary. Under NCLB, state participation became mandatory. The idea is to use NAEP as a “check” on the states’ proficiency levels.

Table 1. Percent of Students in the Nation Labeled “Proficient” or “Advanced” by Three Test Batteries and NAEP

Grade	Reading Comprehension			NAEP 1998	Mathematics			NAEP 2000
	Text X	Test Y	Test Z		Test X	Text Y	Test Z	
1	64	43	14		49	41	12	
2	58	28	41		49	34	40	
3	56	37	14		46	32	5	
4	55	40	24	31	44	34	15	26
5	49	32	33		42	28	30	
6	47	31	18		41	24	12	
7	47	39	25		39	19	21	
8	47	39	33	33	37	23	33	27
9	49	39	11		37	17	9	
10	50	34	15		35	11	12	
11	53	23	25		33	7	24	
12	54	22	29	40	30	5	27	17

Source: H.D. Hoover, “Some Common Misconceptions about Tests and Testing,” 2002 National Council on Measurement in Education Presidential Address, *Educational Measurement: Issues and Practice* (Spring 2003): pp. 5-14, Table 11, p.11.⁵

So, what’s the “real” percentage of eighth-graders who are proficient in math? Test X’s 37 percent, test Y’s 23 percent, test Z’s 33 percent, or NAEP’s 27 percent? The answer is that all these figures are “real” – and equally meaningless without information about the content and skills these tests are measuring, the difficulty of the test items, and where the cut score was set. What’s clear, however, is that all other things being equal, a lower percentage of students would be found proficient, and a higher percentage of schools would fail AYP, in a state that used test Y or Z than in a state that used test X.

Proficient typically means different things in different grades, even on the same test and in the same subject.

Understandably, most people would think that proficient meant the same thing throughout the grades, even though they also know that different, or different levels of, content and skills are taught across the grades. Therefore, if they looked at Test X and Test Y in Table 1 for both reading and math, they would likely conclude that school quality deteriorates from the first to the 12th grade or that students get a lot dumber over time or both. And imagine explaining to parents test Z’s wildly different percentages of proficient students in first, second, third, and other grades!

The fact is, the criterion for proficient (or basic or advanced) is rarely comparable across the grades, even in the same subject. In order for there to be comparability, at least some, if not all, of the following would have to occur: the content and rigor of a state’s academic standards, its tests, and the cut points set on the tests would have to be calibrated across the grades. Few, if any, states can claim that. Therefore, parents and the public are likely, naturally enough, to reach all sorts of invalid conclusions if, say, AYP failure rates are higher in high schools than they are in elementary schools. It just may, or may not, be because the content, tests, and cut points are higher for the high schools.

Parents are in for an especially difficult time because they will get NCLB-mandated school report cards that must break out the percentage of proficient students by grade level. If the results bounce around anything like those on Tests X, Y, and Z – and these are real results from real

⁵ Hoover writes (p.9) that tests X, Y and Z are produced (though not necessarily in the order Hoover lists them) by CTB/McGraw-Hill, Harcourt Educational Measurements, and Riverside Publishing Company.

tests – there’s likely to be a lot of understandable but misdirected anger towards teachers in some grades. As a recent report by the Northwest Evaluation Association (NWEA) put it, “Standards that are not calibrated give students, parents, and educators an inaccurate perception about the child’s standing relative to the expected level of performance. Students are reported as proficient in one grade who may not remain proficient in later grades even if they show normal growth.”⁶

Different percentages of proficient students in different subjects do not necessarily mean that students or schools are stronger or weaker in one subject than the other.

No one would blink an eye if told that students were better in reading than in math, or vice versa. People are not surprised by differences in achievement in different subjects, and few parents have children who are equally strong or weak in all their school subjects. So, in looking back at Table 1, it would seem reasonable to conclude that in virtually every grade and on each of the four displayed tests, a higher percentage of students are proficient in reading than in math. Similarly, if the AYP status report of a certain elementary school showed that 50 percent of the students, on average, were proficient in math and 40 percent were proficient in reading, parents there would almost surely believe that the school was doing a better job in math than in reading.

But they would not necessarily be correct. Is the reading and math content being measured by the reading and math tests of comparable difficulty? Are the reading and math test items of comparable difficulty? How can you even compare achievement in reading and math – or in other subjects – when these subjects are so different?

The fact is you cannot. The best you can do is to compare where the cut points are set on tests of different subjects to see if it is harder to reach proficiency in one than the other. A recent study by the NWEA examined this question of whether proficiency levels are consistent among subject areas, even within the same state. The researchers took the results from reading and math tests given in 14 states and also administered an NWEA test to the same sample of students who had taken the different state tests; the point was to be able to place the results from different state tests onto a common measurement scale so that they could be compared.

As Table 2 shows, in no state were the seventh- or eighth-grade cut scores for proficiency in reading and math set at the same level of difficulty. Moreover, in all states but Oregon, the cut score for math proficiency was more difficult than the standard for reading. In a few states, the difference was quite large. In Arizona, for example, the cut score for proficiency in math would not be reached by 75 percent of NWEA’s norm group of students, while only 47 percent would not reach the proficient cut score in reading.⁷

⁶ G. Gage Kingsbury, Allan Olson, John Cronin, Carl Hauser, and Ron Houser, *The State of State Standards*, Northwest Evaluation Association, <http://www.nwea.org/research/statestudy.html>, November 24, 2003, p. 15.

⁷Ibid., pp. 18-21.

Table 2 – Differences in Reading and Mathematics Percentile Cut Scores for the Proficient Level of Performance – Grade 7 or Grade 8

State	Math cut score (Percentile)	Reading cut score (Percentile)	Difference
Arizona	75	47	28
Colorado	31	12	19
Wyoming	89	74	15
South Carolina	80	68	12
Idaho	46	32	14
Washington	78	67	11
Texas	35	24	11
Minnesota	42	32	10
Illinois	40	32	8
Indiana	42	35	7
California	59	54	5
Montana	36	35	1
Iowa	36	35	1
Oregon	50	58	-8

Source: Kingsbury et al, *The State of State Standards*, Northwest Evaluation Association (NWEA), November 24, 2003, <http://www.nwea.org/research/statestudy.html>, Table 7, p. 19.

Alas, the transparency afforded by the NWEA study is not readily available to parents, the public, and the media – or even to most educators – when they get school, district, state, or national proficiency results. It’s easy to understand, then, why school board members who face tough decisions about how to deploy scarce resources would redirect funds from reading to math when the tests “tell” them that more students are proficient in reading than in math. Maybe that was the appropriate thing to do, and maybe not. What is certain, however, is that you cannot reach a valid conclusion about relative academic strength based only on the percentage of proficient students in one subject versus another.

Proficient means different things in different states.

Since states have different content standards and use different tests and different methods for setting cut scores, obviously the meaning of proficient varies among the states. Under NCLB, states are free to set their own standards for proficiency, which is one reason why AYP school failure rates vary so widely across the states. It’s a lot harder for students to achieve proficiency in a state that has set that standard at a high level than it is in a state that has set it lower. Indeed, even if students in two schools in two different states have exactly the same achievement, one school could find itself on a failed-AYP list simply because it is located in the state whose standard for proficient is higher than the other state’s.

But how great are the differences among the states in their standards for proficiency? The NWEA study also helps to address this question, while underscoring again the differences *within* states in the meaning of proficiency across grades and between math and reading.

As Table 3 shows, states vary, sometimes dramatically, in the rigor of their proficiency standards. For example, the cut score for proficient in grade 4 reading ranged from the 73rd percentile in

Table 3 - Cut Scores Representing "Proficient" or "Meets Standards" Level of Performance on 14 State Assessments
Reading

Grade 3			Grade 4			Grade 5			Grade 6			Grade 7			Grade 8			Grade 9			Grade 10		
State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile
SC	205	67	WY	214	73	SC	220	73	SC	221	63	SC	227	70	WY	232	74	MT	224	43	OR	236	77
CA	200	51	SC	213	70	CA	214	54	CA	216	46	WA	226	67	SC	230	68	IA	224	43	WA	227	51
MN	193	35	WA	207	53	AZ	210	45	MT	211	35	CA	221	50	OR	227	58	ID	221	37	ID	224	44
OR	193	35	CA	205	46	OR	209	42	ID	211	35	MT	218	43	CA	226	54	CO	204	9	MT	224	44
ID	193	35	ID	200	34	IL	207	37	IN	210	32	IA	216	37	AZ	224	49				IA	223	42
MT	193	35	MT	196	26	MT	206	35	IA	209	30	ID	215	35	IN	219	35				CO	209	15
IL	193	35	IA	196	26	ID	206	35	TX	208	28	TX	210	24	MT	219	35				CA	208	14
IN	192	32	CO	191	18	IA	205	32	CO	197	11	CO	206	18	IA	219	35						
IA	191	31				MN	204	30							ID	218	32						
AZ	190	29				TX	204	30							IL	218	32						
TX	179	13				CO	197	18							MN	218	32						
CO	179	13													CO	206	12						

Mathematics

Grade 3			Grade 4			Grade 5			Grade 6			Grade 7			Grade 8			Grade 9			Grade 10		
State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile	State	Cut Score	%ile
SC	208	75	WY	221	83	SC	227	76	SC	235	78	SC	242	78	WY	257	89	MT	242	47	WA	257	73
CA	204	60	WA	218	76	CA	225	70	CA	230	67	WA	242	78	SC	251	80	IA	241	44	MT	247	40
IN	201	50	SC	217	74	AZ	220	59	IN	221	47	CA	238	70	AZ	248	75	ID	240	42	IA	247	40
OR	199	46	CA	212	59	OR	215	46	ID	219	42	ID	225	44	CA	240	59	CO	235	32	OR	245	33
AZ	199	46	ID	205	39	ID	213	41	IA	218	40	MT	224	42	OR	235	50				ID	242	25
MN	198	42	IA	205	39	MT	212	38	MT	218	40	IA	222	38	ID	233	46				CO	233	14
MT	197	39	MT	205	39	IA	212	38	CO	207	19	TX	221	35	MN	231	42				CA	232	13
IA	197	39				MN	210	33				CO	216	26	IN	231	42						
ID	196	36				IL	210	33							IL	230	40						
IL	193	29				TX	209	31							MT	228	36						
						CO	201	15							IA	228	36						
															CO	225	31						

- IN tests students in the fall. Their cut scores were adjusted to reflect equivalent spring performance.
- CO uses the partially proficient level of performance for NCLB reporting. To maintain consistency we report the level each state uses for NCLB reporting here.
- The TX estimate is based on the level for proficient performance that will be implemented in 2005.

Source: Kingsbury et al, *The State of State Standards*, Northwest Evaluation Association, <http://www.nwea.org/research/statestudy.html> on Nov. 24, 2003, Table 4, p. 13.

Wyoming – a very high standard – to the 18th percentile in Colorado.⁸ Similarly, the proficient cut score in grade 10 math ranged from the 73rd percentile in Washington to the 13th percentile in California.

Does this mean that some states’ standards for proficient are “right” and others “wrong”? That is a matter of judgment. What is clear, however, is that the states have very different ideas about what it means to be proficient. Indeed, one of the ways parents can boost their chances of having proficient children is to move to another state! On the other hand, if they want their children to be extremely challenged, they should live in states like Wyoming, Arizona, or South Carolina, where, NWEA tells us, “the standards are set at levels that may be beyond what is needed for a student to have good prospects for completing college.”⁹

NAEP proficiency standards are set very high.

Under NCLB all states must administer NAEP every other year in reading and mathematics in grades 4 and 8, starting in 2003. The idea is to use NAEP as a “check” on states’ assessment results under NCLB or as a benchmark for judging states’ definitions of proficient. If, for example, a state reports a very high percentage of proficient students on its state math test but its performance on math NAEP reveals a low percentage of proficient students, the inference would be that this state has set a relatively easy standard for math proficiency and is trying to “game” NCLB.

But how rigorous are the NAEP standards for proficient?¹⁰ Very rigorous indeed. A few states’ proficiency standards come close, but NAEP’s surpass them all.¹¹

Table 4. Percentile Ranks of the Proficient and the Advanced Levels on the Most Recent NAEP

(Grades 4 & 8 – 2003; Grade 12 Reading – 2002; Grade 12 Math – 2000)

Grade	Proficient		Advanced	
	Reading	Mathematics	Reading	Mathematics
4	69	68	92	96
8	68	71	97	95
12	64	83	95	98

Source: Robert L. Linn, “Assessing Student Achievement: Requirements and Consequences of the No Child Left Behind Act of 2001,” paper presented at the University of California at Davis, Feb. 23, 2004; see also, Robert L. Linn, “Accountability: Responsibility and Reasonable Expectations,” *Educational Researcher* 32, no. 7 (October 2003): pp. 3-13, p. 5, Table 1.

As Table 4 shows, the proficient level on NAEP for grade 4 and 8 reading is set at almost the 70th percentile, while for grade 12, it’s at the slightly less rigorous 64th percentile. The proficiency levels for grade 8 and 12 math are even more challenging. Note, too, the extreme rigor of NAEP’s

⁸ Colorado folded its “partially proficient” standard into its proficient standard for NCLB purposes, which is why its standard is relatively low. The state retains its original proficient standard, which was set quite high, in its own accountability system.

⁹ Kingsbury et al, p. 21.

¹⁰ The lack of comparability of states’ and NAEP’s achievement levels is not the only problem in using NAEP as a “check” on the states. Differences in state and NAEP content standards is another one. This discussion, however, is limited to the former issue.

¹¹ D. McLaughlin and V. Bandeira de Mello, “Comparison of State Elementary School Mathematics Achievement Standards Using NAEP 2000,” paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA, April 2002.

advanced levels in reading and math in grades 4, 8 and 12, which range from the 92nd to the 98th percentile!

It would not be unreasonable to think that the proficiency levels on NAEP represent a standard of achievement that is more commonly associated with fairly advanced students. But if NAEP is NCLB's "check" on the states, then NAEP proficiency levels are what *all* fourth-, eighth-, and 12th-graders are expected to achieve by 2014.

Can it be done? The percentage of fourth-, eighth-, and 12th-grade students who scored at proficient or above on NAEP increased throughout the 1990s, which is good news. However, if past rates of progress on NAEP are a guide, it will not take 12 years to achieve 100 percent proficiency, as NCLB requires. Rather, in math, it will take 57 years for all fourth-graders to reach that goal, 61 years for eighth-graders, and 166 years for 12th-graders. It will take even longer to attain 100 percent proficiency in reading because, even though the percentage of proficient students is higher in reading than in math, the rate of progress on reading NAEP has been somewhat slower.¹²

Is an acceleration of past rates of progress on NAEP achievable? Likely so. But is it possible to progress four to twelve times faster than we have in math and even more in reading? As one of the world's most respected testing and measurement experts concluded, "Such rapid acceleration would be nothing short of miraculous."¹³

NCLB's requirement of 100 percent proficiency is not the same thing as requiring that all students achieve at grade level.

As we have seen, proficiency does not have a single meaning. It essentially represents judgments about where to set cut scores on tests. They can be set high or low or somewhere in between, depending on the standard-setting method used and the views of the standard setters.

Grade level is an entirely different concept. It does not represent a judgment about what scores a student *ought* to achieve. Rather, it is an "is" statement: a description of how a student achieved in relation to "typical" students who took the same test in the same grade at the same time of year (the norm population, or the representative sample of students who took the test). So, while proficiency under NCLB is subjective, a judgment call, grade level is objective in the sense that it is based on the actual performance of students. You cannot, for example, decide that being on grade level in 12th-grade math should represent an Einstein-like or, at the other extreme, pre-numerate level of achievement on a particular test if the average level of performance on that test was neither Einstein-like or pre-numerate.¹⁴

So, let's say Jane's parents got a report that their fifth-grade daughter is on grade level in Reading Comprehension on the Iowa Test of Basic Skills (ITBS); Jane took the test in the spring of fifth grade and scored 5.4. What does that mean? According to ITBS's recommended definition, it means that Jane is on grade level because her grade equivalent score of 5.4 is in the range of 4.2-7.9, which corresponds to the performance of students scoring in the "average" range during the

¹² Linn (2003), pp. 5-7.

¹³ Ibid., p. 6.

¹⁴ Achievement standards, like proficiency standards, are increasingly being set on norm-referenced tests, a process that transforms such tests into criterion-referenced tests. This discussion is confined to traditional means of determining grade level.

spring months of grade 5.¹⁵ And Jane's parents would very likely understand this because the definition is akin to the one they hear when their pediatrician tells them that their 54-inch, 10-year-old is "average" in height because her height falls within the range of 50-59 inches that is typical or "normal" for 10-year-old girls.¹⁶

Now, contrast this with the report Jane's parents must get under NCLB. Jane scored 54 on the state's reading test and did not achieve the cut score for proficient, which was 59. Does that mean that Jane is below grade level in reading? Her parents have no way of knowing. Perhaps they'll get some description of what a student who is proficient in reading is supposed to know and be able to do. They'll likely learn the percentage of reading-proficient students in Jane's grade and school (if other grades were tested) or district or even in the state. But that still won't tell them whether or not Jane is at grade level. It would be like their pediatrician saying that, at 54 inches, Jane is short because she is not above 59 inches (which represents only five percent of 10-year-old girls); that still doesn't tell her parents whether Jane's height is normal.

But because Jane is not proficient in reading, her parents will no doubt be concerned. Jane's classmate, John, did score at the proficient level on the state's reading test, but his parents are likely to be concerned, too, because the school is on a list of schools that failed to make "adequate yearly progress" (AYP) under NCLB. Although 55 percent of its students were at the proficient level in reading and 50 percent were proficient in math, the AYP target set for every school in that state required that 45 percent of the school's students as a whole and in each of the school's subgroups¹⁷ be at the proficient level in reading and 40 percent be at the proficient level in math that year.

So why did John and Jane's school fail AYP? The special education subgroup was 15 points below the state's targets in both subjects, while the low-income student subgroup, which included Jane, was five points below. So even though the school as a whole and each of its subgroups made above-average progress from the previous year – a commendable accomplishment – the school failed AYP because it didn't precisely hit each of its AYP targets.

But does that mean that 45 percent of the school's students as a whole are below grade level in reading and 50 percent are below grade level in math, while even heftier percentages of subgroup students fall below grade level? Not necessarily. In fact, even if all students in the school were on grade level, the school could still fail to make AYP. Students' proficiency scores and their grade levels are not the same thing.

Pointing out that proficiency and grade level are different is not an argument for the superiority of grade-level reporting. The range of achievement that constitutes grade-level achievement is so great that it is hardly a standard at all.¹⁸ Indeed, there are excellent arguments for setting

¹⁵ H.D. Hoover et al, *The Iowa Tests: Interpretive Guide for School Administrators*, University of Iowa, 2003 (Riverside Publishing), p. 67. The authors also note that many people define "average" as the 50th percentile or, in the case of this example, 5.8, which, in contrast to Jane's 5.4, is the exact average for students who took that test in the spring of fifth grade. But, they say, this is "a very narrow view...that does not take error and student variability into account. A more common and useful definition of average is performance between the 25th and 75th percentiles. This range....discourages the overinterpretation of small differences in student performances" (p. 68).

¹⁶ <http://sickkid.net/charts/pdf.charts/g-statage-wtage-e.pdf>.

¹⁷ NCLB designates student subgroups by race/ethnicity; poverty; English language learners; disability; and migrant status.

¹⁸ See, for example, "How Achievement Varies," *ETS Policy Notes*, Educational Testing Service, Princeton, NJ, Summer 1993, p. 1, 5-6.

standards/cut scores on the basis of ambitious, but *reasonable*, judgments about where student performance *ought* to be rather than according to how students actually perform. We certainly do it all the time when it comes to things like, say, weight, which, unlike achievement, we are frequently urged to reduce because Americans' actual weight is far above "ideal" weights.

But weight-reduction goals, unlike the goal of NCLB, are transparent. We all know, for example, what a "pound" is; its meaning and measurement do not vary across the states.¹⁹ And unlike the many meanings of academic proficiency, we can readily find the standards and ranges for "ideal" weights. We also know that extremely overweight people have a lot more pounds to lose to achieve their ideal weights, so it will take them longer and require more support than those who are merely a bit overweight. Yet NCLB does not similarly recognize the different academic starting points of students, including those whose achievement lags even before they start school; the interim achievement targets are the same for all students, whether they are well below or well above them.

And so, unlike the praise we lavish on people who have lost a lot of weight in, say, a year, even though they've fallen short of their targets, when schools fall short of AYP – no matter how unreasonable their targets and no matter how great their progress – they will be put on a failing list. The only comfort they will have is cold comfort indeed: like overweight Americans, they will be in the majority. Indeed, if the experts are right, in a couple of years, virtually every public school district in the nation will be on a failed-AYP list, and by 2014, the year 100 percent of our students are supposed to be proficient, almost every school will have failed to make AYP for one or more years. The reason isn't to be found in the near-universal deficiency of our public schools – though there are certainly failures among them – but in the deficiencies of the law.

American Federation of Teachers, AFT-CIO
555 New Jersey Ave. N.W. • Washington, DC 20001 • 202/879-4400 • www.aft.org

¹⁹ This is not an argument either for national standards or national tests, but, rather, an analogy meant simply to illustrate the relative lack of public understanding about NCLB proficiency standards.